

Impact of the channel state information on the energy-efficiency of MIMO communications

Fernando Rosas¹ and Christian Oberli²

¹ Departement Elektrotechniek, Katholieke Universiteit Leuven

² Departamento de Ingeniería Eléctrica, Pontificia Universidad Católica de Chile

Abstract

Multiple-input multiple-output (MIMO) techniques provide attractive tools for reducing the energy consumption of wireless communications. Even though some research has been reported on this topic, the ground rules and trade-offs by which the MIMO physical layer parameters should be chosen in order to achieve energy efficiency have not yet been formally established. In this paper, we study the impact of having transmitter-side channel state information (TCSI) by analyzing the energy consumption of the singular value decomposition (SVD), Beamforming, Zero Forcing and Generalized Alamouti MIMO schemes.

We show that MIMO schemes with large diversity degree are energy-optimal for transmitting data over long transmission distances, while schemes with large multiplexing gain are more energy-efficient for performing short-range communications. We further show that large antenna arrays are optimal for performing long-range transmissions, however they are suboptimal for transmitting over short link distances. Finally, our results suggest that although exploiting TCSI provides important savings in long-range communications, it is not critical for transmissions over short link distances.

Corresponding author:

Fernando Rosas, fernando.rosas@esat.kuleuven.be.

Impact of the channel state information on the energy-efficiency of MIMO communications

I. INTRODUCTION

Multiple-input multiple-output (MIMO) communication systems were originally introduced as a way for boosting data rates or for improving the reliability of wireless links [1], [2]. Recently, researchers started to realize that the MIMO techniques could also be used for reducing the energy consumption of wireless communications (see e.g. [3]–[15]). Despite the recent interest, the ground rules and trade-offs by which the main MIMO physical layer parameters such as modulation scheme, constellation size and irradiation power should be chosen in order to achieve energy efficiency have not yet been formally established [16].

For taking full advantage of a MIMO link in terms of the multiplexing-diversity trade-off, it is valuable to have transmitter-side channel state information (TCSI). In effect, several MIMO modulations require to distribute the transmitted power only among the most favorable angular domains [17]. Moreover, the capacity of a MIMO communication link is reduced significantly when there is no TCSI available [2], [18].

While the effect of the TCSI on the MIMO channel capacity is well known, the question about its effect on the achievable energy-efficiency has not been reported yet in the available literature. In this paper, we give an answer to this question by presenting an energy consumption model of MIMO communications under fading channels, which is an extension of the one introduced in [19]. Most of other energy consumption models reported so far in the literature [6]–[14] are based on the notion of capacity of a MIMO random fading channel, which plays a key role in linking the rate of information transfer, the signal-to-noise ratio (SNR) and the energy consumption. Our approach avoids the concept of channel capacity, sharing some features with the one reported in [3] but finding a more straightforward mathematical formulation that enables a more insightful analysis. Moreover, our MIMO energy consumption model is also novel because it considers the effect of the retransmissions required to guarantee error free transmissions.

Using our model, we analyze the performance of the MIMO modulations shown in Table I. We show that modulations with large diversity gain are best for achieving energy-efficient

transmissions over long link distances. On the other hand, modulations with large multiplexing gain are optimal for short-range communications. This result generalizes the findings presented in [20] on the optimal constellation size for single antenna systems.

TABLE I
MIMO MODULATIONS

	TCSI	No TCSI
Multiplexing	MIMO SVD	Zero-Forcing
Diversity	MIMO Beamforming	Generalized Alamouti

When studying the impact of the antenna array size on the energy consumption, we show that devices equipped with a small number of antennas are more energy-efficient than large antenna arrays for performing short-range transmissions. Conversely, large antenna arrays can achieve significant savings when the transmission is done over long distances.

Finally, we show that for long-range communications the energy-consumption of the studied MIMO schemes that do not exploit TCSI can be up to 100 times larger than the consumption of studied schemes that do use TCSI. The relative losses become less significant as the transmission distance shortens, and are negligible for very short link distances (less than 20 meters for typical low power devices). This suggest that having TCSI is only critical from an energy point of view when the link distance is large.

The rest of this paper is organized as follows: Section II presents the energy consumption model, Section III specifies the dependence of the energy consumption on the SNR and Section IV presents an analysis of the energy consumption of MIMO SVD, Zero Forcing and Generalized Alamouti schemes for various link distances and antenna array sizes. Finally, Section V presents our conclusions.

II. ENERGY CONSUMPTION MODEL

Our goal is to determine the total energy that is necessary for transferring one bit of data successfully in a point-to-point packet-switched MIMO communication. We call such a bit a *goodbit* [8]. Following [20], it is assumed that every frame transmitted in the *forward* direction

is matched by a feedback frame in the *reverse* direction that acknowledges correct reception or requests a retransmission. It is also assumed that the radiated power is determined based on knowledge of the statistics of the SNR at the decision stage of the receiver. It is further assumed that all frames in both directions are always detected and that all feedback frames are decoded without error.

Our interest is to compare the energy consumption of different MIMO schemes. For this, the choice of channel-coding method of the source data is irrelevant, reason for which we leave channel coding out of this work.

In the sequel, Section II-A presents the analysis of the energy consumption of a transceiver equipped with N_t antennas that transmits one payload frame and receives the corresponding feedback frame. The reverse case —a transceiver equipped with N_r antennas that receives one payload frame and transmits the corresponding feedback frame— follows by analogy. Section II-B then synthesizes the total energy consumption model. The energy consumption analysis has been made for the MIMO transceiver architecture shown in Figure 1, which is frequently used among academic and commercial products [21]–[23].

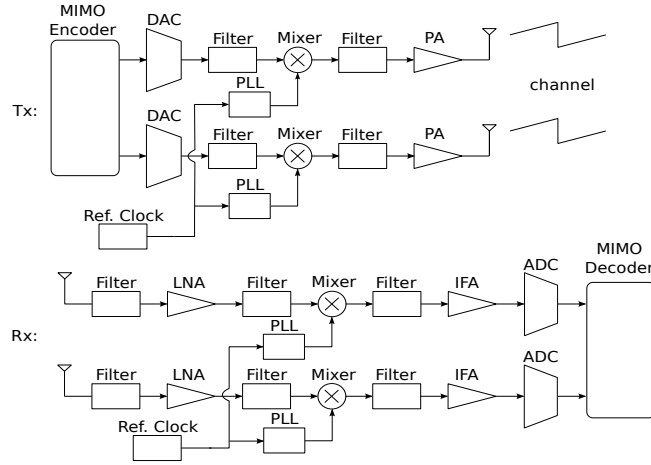


Fig. 1. Common architecture of a MIMO transceiver.

Notation: in the following, symbols with a hat (i.e. \hat{A}) will denote quantities that account the whole MIMO transceiver, in contrast to quantities related to individual transceiver branches which shall be represented without a hat.

A. Components of energy consumption of the forward transceiver

The energy consumption of the MIMO transceiver that transmits forward frames and receives feedback frames is composed of six terms, each one described next.

1) *Startup energy consumption:* It is assumed that the transmitter is by default in a low power consumption (sleep) mode. Hence, it must be brought online before it can make a transmission. Let us denote $\hat{\mathcal{E}}_{\text{st,tx}}$ as the total startup energy divided by the number of data bits that will be transmitted before the transceiver goes into low power consumption mode again. In a MIMO system, $\hat{\mathcal{E}}_{\text{st,tx}}$ is largely dominated by the energy spent in the stabilization of the N_t phase-lock-loops (PLL) of the transceiver (see Figure 1), while startup costs of components common to all branches are negligible [24]. Therefore $\hat{\mathcal{E}}_{\text{st,tx}} = N_t \mathcal{E}_{\text{st}}$, where \mathcal{E}_{st} is the startup energy consumption per branch per goodbit.

2) *RF electronic consumption:* The total air time per forward frame is composed by T_L seconds used for the transmission of the L payload bits that compose the frame, T_H seconds for the transmission of the H bits that compose the frame header (which carries addresses and parameters) and T_O seconds used for the transmission of overhead signals for tasks such as acquisition, channel estimation and synchronization [25]. The air time per payload bit is therefore

$$\hat{T}_b = \frac{T_L + T_H + T_O}{L} . \quad (1)$$

Let us define R_s as the physical layer symbol-rate of each transmitter antenna branch after the MIMO encoding, and ω as the multiplexing gain of the MIMO modulation scheme. Therefore, ωR_s is the source symbol rate feeding the MIMO encoder. Source symbols are modulated with an M -ary modulation, hence

$$R_b = \omega b R_s \quad (2)$$

is the total (source) bit rate of the MIMO system, where $b = \log_2 M$. In the case that ω is an integer and constellations of sizes M_1, \dots, M_ω are used simultaneously (such as in the case of the MIMO SVD scheme), then b is defined as the average $b = \omega^{-1} \sum_{k=1}^{\omega} \log_2 M_k$. By considering that header bits are sent using a binary modulation, and noting that $L/T_L = R_b$, then one can express \hat{T}_b as

$$\hat{T}_b = \frac{1}{R_s} \left(\frac{1}{\omega b} + \frac{H}{\omega L} + \frac{N_t O_a + O_b}{L} \right) , \quad (3)$$

where O_a is the acquisition overhead per branch and O_b is the remaining overhead, which is approximately independent of the antenna array size. Both O_a and O_b are measured in bits.

The RF electronic components involved in the transmission processing consume $\hat{P}_{\text{el,tx}}$ Watts, which is largely dominated by the consumption of passband processing components such as filters, mixers and frequency synthesizers engaged in the forward transmission [26]. Hence, it can be inferred from Figure 1 that $\hat{P}_{\text{el,tx}}$ grows linearly with the number of transmit branches. Therefore, the energy consumed in the transmission processing per payload bit can be expressed as

$$\hat{\mathcal{E}}_{\text{el,tx}} = \hat{P}_{\text{el,tx}} \hat{T}_b = N_t P_{\text{el,tx}} \hat{T}_b, \quad (4)$$

where $P_{\text{el,tx}}$ stands for the electric power consumed by each branch of the transmitter.

3) *Energy consumption due to electromagnetic radiation:* Each frame is aired out of all N_t branches of the transceiver. It is defined $P_A^{(j)}$ as the power radiated by the antenna of the j -th branch, which is supplied by a corresponding power amplifier (PA) (Figure 1). The power consumption of the j -th PA, $P_{\text{PA}}^{(j)}$, is modeled as

$$P_{\text{PA}}^{(j)} = \frac{\xi}{\eta} P_A^{(j)}, \quad (5)$$

where ξ is the peak-to-average ratio of the transmitted signal and η is the drain efficiency of the PA [26]. Thus, the energy per payload bit used for electromagnetic radiation is given by

$$\hat{\mathcal{E}}_{\text{RF}} = \left(\sum_{j=1}^{N_t} P_{\text{PA}}^{(j)} \right) \hat{T}_b = \hat{P}_{\text{PA}} \hat{T}_b, \quad (6)$$

where \hat{T}_b is given by (1), and \hat{P}_{PA} has been defined as a shorthand notation for the total power consumption of all the PAs.

4) *Energy consumption of electronic components due to the processing of feedback frames:* Feedback frames are assumed to last $F/(\omega b R_s)$ seconds, where F is the number of bits that compose the feedback frame and $\omega b R_s$ gives the total bit-rate (c.f. (2)). During this time, the MIMO receiver consumes $\hat{P}_{\text{el,rx}}$ Watts, which mainly includes the power needed for energizing the passband receiver elements (low-noise amplifiers, mixers, filters, frequency synthesizers, etc.) of all the branches [26]. Hence, $\hat{P}_{\text{el,rx}}$ grows linearly with the number of transceiver branches (N_t). Therefore, the energy per forward payload bit spent by the transmitter for decoding the corresponding feedback frame is

$$\hat{\mathcal{E}}_{\text{fb,rx}} = \hat{P}_{\text{el,rx}} \frac{F}{\omega b R_s L} = N_t P_{\text{el,rx}} \hat{T}_{\text{fb}}, \quad (7)$$

where $P_{\text{el,rx}}$ is the electronic power consumption of one receiver branch and

$$\hat{T}_{\text{fb}} = F/(\omega b R_s L) \quad (8)$$

is the feedback time taken per payload bit.

5) *Baseband electronic consumption:* The computations required for calculating the MIMO encoder are jointly the most demanding baseband operations that the transceiver performs. Each complete computation involves a set of K different arithmetic operations, each of which has an energy consumption \mathcal{E}_k and is performed $n_k^{(\text{tx})}$ times during the encoder computation. Thus, the energy consumption of the transmitter for computing the MIMO encoder, $\hat{\mathcal{E}}_{\text{cod,tx}}$, is given by

$$\hat{\mathcal{E}}_{\text{cod,tx}} = \sum_{k=1}^K \mathcal{E}_k n_k^{(\text{tx})} . \quad (9)$$

If the operations are performed by an arithmetic processing unit (APU), the energy consumption of the k -th operation can be modeled as [27]

$$\mathcal{E}_k = V_{\text{dd}} I_0 \Delta t_k , \quad (10)$$

where V_{dd} is the APU operating voltage and I_0 is the average current during the execution time of the arithmetic operations. It is to be noted that I_0 depends on V_{dd} and on the APU's clocking frequency, f_{APU} . Δt_k is the time required for executing the k -th operation, which is related to f_{APU} and to the number of clock cycles required by the operation, c_k , as

$$\Delta t_k = \frac{c_k}{f_{\text{APU}}} . \quad (11)$$

Replacing these terms in (9), the mean energy required for computing the MIMO encoder in the transmitter is given by

$$\hat{\mathcal{E}}_{\text{cod,tx}} = \frac{V_{\text{dd}} I_0}{f_{\text{APU}}} \sum_{k=1}^K c_k n_k^{(\text{tx})} . \quad (12)$$

The MIMO encoder has to be recalculated each time the MIMO channel has changed significantly. The rate of change of the channel is measured by its coherence time, which is defined as [28]

$$T_c = \frac{9v_0}{16\pi v_m f_c} , \quad (13)$$

where v_0 is the speed of light, f_c is the carrier frequency and v_m is the maximum speed found in the mobile environment. The energy consumption of calculating one MIMO encoder is shared

among all the payload bits that are transmitted within one coherence time. That number of bits can be approximated by $T_c/(\hat{T}_b + \hat{T}_{fb})$, where \hat{T}_b and \hat{T}_{fb} are as defined in (1) and (8) respectively*. Therefore, the mean energy consumption per payload bit can be expressed as

$$\hat{\mathcal{E}}_{bb,tx} = \frac{\hat{T}_b + \hat{T}_{fb}}{T_c} \hat{\mathcal{E}}_{cod,tx} . \quad (14)$$

B. Total energy per goodbit

The material presented in Section II-A allows for stating our model of the total energy consumption. Concretely, the energy consumed per goodbit by the transmitter of forward frames, which also decodes feedback frames, is given by

$$\hat{\mathcal{E}}_T = \hat{\mathcal{E}}_{st,tx} + (\hat{\mathcal{E}}_{bb,tx} + \hat{\mathcal{E}}_{el,tx} + \hat{\mathcal{E}}_{RF} + \hat{\mathcal{E}}_{fb,rx})\tau \quad (15)$$

$$= \hat{\mathcal{E}}_{st,tx} + \left[\frac{\hat{T}_b + \hat{T}_{fb}}{T_c} \hat{\mathcal{E}}_{cod,tx} + (\hat{P}_{el,tx} + \hat{P}_{PA})\hat{T}_b + \hat{P}_{el,rx}\hat{T}_{fb} \right] \tau . \quad (16)$$

Above, τ denotes the number of transmission trials until a forward frame is decoded without error. This random variable, whose distribution will be studied in III-B, is introduced for accounting the losses introduced by retransmissions due to forward frames that get decoded with errors at the receiver.

By analogy, the total energy used by the receiver for demodulating τ forward transmissions and for transmitting the corresponding τ feedback frames is

$$\hat{\mathcal{E}}_R = \hat{\mathcal{E}}_{st,rx} + \left[\frac{\hat{T}_b + \hat{T}_{fb}}{T_c} \hat{\mathcal{E}}_{cod,rx} + \hat{P}_{el,rx}\hat{T}_b + (\hat{P}_{el,tx} + \hat{P}_{PA})\hat{T}_{fb} \right] \tau . \quad (17)$$

Above, it has been introduced $\hat{\mathcal{E}}_{st,rx} = N_r \mathcal{E}_{st}$ as the startup energy consumption of the receiver (c.f. Section II-A1), and $\hat{\mathcal{E}}_{cod,rx}$ as the energy consumption of calculating a MIMO decoder in the receiver, which is defined using (9) with obvious modifications.

The total energy consumption per goodbit, which is given by the sum of (15) and (17), can be written as

$$\hat{\mathcal{E}}_b = \hat{S} + \left[\hat{B} + \hat{P}_{el} + \hat{P}_{PA} \right] \hat{T} \tau , \quad (18)$$

where it has been defined $\hat{S} = \hat{\mathcal{E}}_{st,tx} + \hat{\mathcal{E}}_{st,rx}$ as the total startup energy consumption per goodbit, $\hat{B} = (\hat{\mathcal{E}}_{cod,tx} + \hat{\mathcal{E}}_{cod,rx})/T_c$ as the total baseband processing power consumption, $\hat{P}_{el} = (\hat{P}_{el,tx} +$

*The proposed approximation holds as long as there are enough data bits to transmit continuously during one coherence time.

$\hat{P}_{\text{el,rx}}$) as the total power consumed by electronic components and $\hat{T} = \hat{T}_{\text{b}} + \hat{T}_{\text{fb}}$ the total time per payload bit per transmission trial.

It is crucial to note that, because of τ , $\hat{\mathcal{E}}_{\text{b}}$ is a random variable whose distribution depends on the antenna array sizes, frame size, modulation type and on the received SNR during the t -th transmission trial. Its mean value is

$$\bar{\mathcal{E}}_{\text{b}} = \mathbb{E} \left\{ \hat{\mathcal{E}}_{\text{b}} \right\} = \hat{S} + \left[\hat{B} + \hat{P}_{\text{el}} + \hat{P}_{\text{PA}} \right] \hat{T} \bar{\tau} . \quad (19)$$

Expressions for $\bar{\tau}$ will be studied in III-B.

III. ENERGY CONSUMPTION AS A FUNCTION OF THE SNR

We seek to understand the relationship between the SNR at the decision stage of the receiver and the energy consumption per goodbit in a MIMO system. To achieve this, we analyze the PAs total power consumption (\hat{P}_{PA}) and the mean number of transmission trials ($\bar{\tau}$), both of which are functions of the received SNR.

A. Total power consumption of the PA as a function of the SNR

First, the case in which there is no TCSI is addressed, and afterwards the case of MIMO SVD is considered.

1) *No TCSI case:* In general, the mean power radiated by the j -th antenna, $\bar{P}_{\text{A}}^{(j)}$, attenuates over the air with path loss and reaches at the receiver with a mean power given by

$$\bar{P}_{\text{rx}}^{(j)} = \frac{\bar{P}_{\text{A}}^{(j)}}{A_0 d^\alpha} , \quad (20)$$

where A_0 is a parameter that depends on the transmitter and receiver antenna gains and on the transmission wavelength, d is the distance between transmitter and receiver and α is the path loss exponent [28].

It has been shown that if the transmitter has no knowledge of the MIMO channel, then no consistent benefits can be achieved by using unequal power allocation between the transmitted symbols [17]. Therefore, we can assume that all the transmitted symbols are radiated with the same mean power. If σ_{s}^2 is the average received power per symbol at the input point of the decision stage of the receiver (which is located after the MIMO decoder), the total received signal power is given by

$$\sum_{j=1}^N \bar{P}_{\text{rx}}^{(j)} = \omega \sigma_{\text{s}}^2 , \quad (21)$$

with ω as defined in Section II-A5. Let us denote the noise power at the decision stage by σ_n^2 and the SNR as $\bar{\gamma} = \sigma_s^2 / \sigma_n^2$. By considering (5), (20) and (21), the total power consumption of the PAs (c.f. Section II) takes the form

$$\hat{P}_{\text{PA}} = \frac{\xi}{\eta} \sum_{j=1}^N \bar{P}_A^{(j)} = \frac{\xi A_0 d^\alpha}{\eta} \sum_{j=1}^N \bar{P}_{\text{rx}}^{(j)} \quad (22)$$

$$= \frac{\xi A_0 d^\alpha \omega \sigma_n^2}{\eta} \bar{\gamma} = \omega A d^\alpha \bar{\gamma} , \quad (23)$$

with A a constant. In general, $\sigma_n^2 = N_0 W N_f M_L$, where N_0 is the power spectral density of the baseband-equivalent additive white Gaussian noise (AWGN), W is the transmission bandwidth, N_f is the noise figure of the receiver's front end and M_L is a link margin term which represents any other additive noise or interference [3].

2) *MIMO SVD*: In this case, the transmitter sends symbols using the n most favorable angular directions of the MIMO channel, which are provided by the first n right-singular vectors of the channel matrix \mathbf{H} , which are associated with the largest singular values [17]. These directions create n non-interfering parallel channels (*eigenchannels* in the following). It is direct to see that in this case $\omega = n$.

Let us denote $\bar{P}_{\text{tx,eig}}^{(k)}$ the power that is allocated by the transmitter to the k -th eigenchannel, where $k \in \{1, \dots, \omega\}$. The irradiation per transceiver branch is given by

$$\bar{P}_A^{(j)} = \sum_{k=1}^{\omega} |v_{j,k}|^2 \bar{P}_{\text{tx,eig}}^{(k)} , \quad (24)$$

where $\bar{P}_A^{(j)}$ is as defined in (II-A3) with $j \in \{1, \dots, N\}$ and $v_{j,k}$ is the coefficient of the j -th row and k -th column of the right unitary matrix \mathbf{V} which is obtained from the singular value decomposition $\mathbf{H} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\dagger$, where $(\cdot)^\dagger$ stands for the hermitian operator. Even though the eigenchannels are not physical antennas, they satisfy a relationship similar to (20):

$$\bar{P}_{\text{rx,eig}}^{(k)} = \frac{\bar{P}_{\text{tx,eig}}^{(k)}}{A_0 d^\alpha} , \quad (25)$$

where $\bar{P}_{\text{rx,eig}}^{(k)}$ is the received power from the corresponding eigenchannel that is obtained after the MIMO SVD decoder. Note that the SNR at the decision stage of the k -th eigenchannel can be written as

$$\text{SNR}_k = \frac{\lambda_k \bar{P}_{\text{rx,eig}}^{(k)}}{\sigma_n^2} = \lambda_k \phi_k \omega \bar{\gamma} , \quad (26)$$

where λ_k is the square of the k -th singular value of the channel matrix \mathbf{H} [29], $\phi_k = \bar{P}_{\text{rx,eig}}^{(k)} / (\sum_{k=1}^{\omega} \bar{P}_{\text{rx,eig}}^{(k)})$ is the proportion of the total power that was allocated to the k -th eigenchannel and $\bar{\gamma} = \omega^{-1} \left(\sum_{k=1}^{\omega} \bar{P}_{\text{rx,eig}}^{(k)} / \sigma_n^2 \right)$ is the average SNR among the used eigenchannels at the decision stage.

Using the definition of \hat{P}_{PA} (cf. Section II-A3), (5), (24), (25), (26) and considering the definition of $\bar{\gamma}$ given above, then the total power consumption of the PAs can be written as

$$\hat{P}_{\text{PA}} = \frac{\eta}{\xi} \sum_{j=1}^N \sum_{k=1}^{\omega} |v_{j,k}|^2 \bar{P}_{\text{tx,eig}}^{(k)} = \frac{\eta}{\xi} \sum_{k=1}^{\omega} \bar{P}_{\text{tx,eig}}^{(k)} \quad (27)$$

$$= \frac{\xi A_0 d^\alpha}{\eta} \sum_{k=1}^{\omega} \bar{P}_{\text{rx,eig}}^{(k)} = \frac{\xi A_0 d^\alpha}{\eta} \omega \sigma_n^2 \bar{\gamma} = \omega A d^\alpha \bar{\gamma} , \quad (28)$$

reaching the same result than in III-A1 in a different context. Above, the fact that $\sum_{j=1}^N |v_{j,k}|^2 = 1$ is used, which is a consequence of \mathbf{V} being a unitary matrix.

B. $\bar{\tau}$ as function of the SNR

The distribution of the number of transmission trials until a frame is correctly decoded (τ , c.f. Section II-B) depends on the probability of decoding the frame with error during the t -th transmission trial, which will be denoted as $P_f(t)$. In general, $P_f(t)$ are also random variables (specifically, they are conditional probabilities in a measure-theoretic sense [30]), whose distributions depend on the antenna array sizes, frame size, modulation type and on the received SNR during the t -th transmission trial.

The distribution of the number of transmission trials until a frame is correctly decoded (c.f. Section II-B) can be found explicitly in the case of fast-fading channels. Consider the assumption that the probabilities of frame error of each transmission trial are a set of i.i.d. random variables. Define their mean value as $\bar{P}_f = \mathbb{E}\{P_f(t)\}$, where the index t has been dropped for simplicity of notation. Then, it is proved in the Appendix A that

$$\mathbb{P}\{\tau = r\} = (1 - \bar{P}_f) \bar{P}_f^{r-1} , \quad (29)$$

which implies that τ distributes as a Geometric random variable with parameter $1 - \bar{P}_f$. It is interesting to note that this result holds for any antenna array size or MIMO modulation scheme. In fact, although these aspects affect the distribution of P_f , τ only depends on its mean value, being independent of its variance or higher moments.

Using well known properties of geometric random variables, it can be seen that

$$\bar{\tau} = \frac{1}{1 - \bar{P}_f} . \quad (30)$$

This is consistent with previous results for single antenna systems [20]. In the following, expressions for \bar{P}_f under different MIMO modulations are derived. In general Rayleigh fading statistics will be considered, as it represents the most unfavorable scenario.

1) *No TCSI case*: Let us say that the H bits of header are transmitted using a binary modulation for minimizing their probability of error, and the L bits of payload are modulated using an uncoded M -ary modulation, with $b = \log_2 M$ bits per symbol. Hence, the frame is composed by H binary symbols and L/b M -ary symbols. It is assumed that the transmitter is equipped with a deep interleaver [31], which completely decorrelates a time-varying MIMO channel between any successive symbols. Therefore, the mean frame error rate \bar{P}_f can be written in terms of mean symbol error rate $\bar{P}_M(\bar{\gamma})$ and the mean binary modulation symbol error rate $\bar{P}_{\text{bin}}(\bar{\gamma})$ for a given channel statistics as

$$\bar{P}_f = 1 - [1 - \bar{P}_{\text{bin}}(\bar{\gamma})]^H [1 - \bar{P}_M(\bar{\gamma})]^{L/b} . \quad (31)$$

It can be shown [32], [33] that, under Rayleigh fading statistics, the symbol error rate (SER) of Zero-Forcing and Generalized Alamouti modulations over narrowband MIMO channels are equal to the SER of a single-antenna Nakagami- m fading channel with appropriate parameters (for completeness, the derivations are provided in Appendix B). The SNR γ of a Nakagami- m fading channel distributes as a Gamma random variable $\Gamma(m, \mu)$, whose p.d.f. is given by [34]

$$f_{\mu,m}(\gamma) = \left(\frac{m}{\mu}\right)^m \frac{\gamma^{m-1}}{\Gamma(m)} \exp\left(-\frac{m\gamma}{\mu}\right) . \quad (32)$$

Above, μ is the mean power gain and m is the diversity gain of the channel [29]. Hence, using (32) and the results presented in Appendices B-2 and B-3, the SER of Zero-Forcing and Generalized Alamouti MIMO modulation schemes can be calculated as

$$\bar{P}_M(\mu\bar{\gamma}, m) = \int_0^\infty P_M(\gamma) f_{\mu,m}(\gamma) d\gamma , \quad (33)$$

with parameters μ and m given in Table II. It is useful to have in mind that $\bar{P}_M(\mu\bar{\gamma}, m) \propto (\mu\bar{\gamma})^{-m}$ in the large SNR regime [31]. The integral of the right hand side of (33) can be evaluated using the techniques presented in [34].

TABLE II
PARAMETER VALUES OF THE SER OF VARIOUS MIMO MODULATION SCHEMES

MIMO modulation	Power Gain (μ)	Diversity gain (m)
Zero-Forcing	1	$N_r - N_t + 1$
Generalized Alamouti	N_t	$N_t N_r$

Finally, using (30), (31) and (33) it is found that

$$\bar{\tau}^{N_t, N_r}(\bar{\gamma}) = [1 - \bar{P}_{\text{bin}}(\mu\bar{\gamma}, m)]^{-H} [1 - \bar{P}_M(\mu\bar{\gamma}, m)]^{-L/b}, \quad (34)$$

where μ and m have to be replaced with adequate values following Table II.

2) *MIMO SVD*: There are different ways in which the bits that compose a frame can be fed into the SVD encoder. In particular, it has been shown that a pseudo-random feeding of the encoder outperforms an ordered feeding [19]. Hence, we will consider the case in which the bits of each frame are assigned to the ω used eigenchannels following a different order for each transmission trial in a pseudo-random fashion. If the ω eigenchannels are equipped with constellations of sizes M_1, \dots, M_ω and $b = \sum_{k=1}^{\omega} \log_2 M_k$, then the frame will be composed of approximately L/b symbols. Because of the law of the large numbers this approximation is tight when $L \gg \max_{k=1, \dots, \omega} \log_2 M_k$ holds, which happens in many cases of interest (in the following we will use it as an equality for simplicity).

Let us define E_p as the event that a payload symbol was decoded with error. Appendix B-1 in conjunction with (26) shows that, for a given realization of the MIMO channel matrix \mathbf{H} , the average symbol error rate of the pseudo-random feeding is given by

$$P\{E_p | \mathbf{H}\} = \frac{1}{\omega} \sum_{k=1}^{\omega} P_{M_k}(\lambda_k \phi_k \omega \bar{\gamma}) \quad (35)$$

where λ_k is the k -th eigenvalue of the matrix \mathbf{H} (see Appendix B) and ϕ_k is as defined in Section III-A2. In a similar fashion, by defining E_h as the event that a header symbol is decoded with error, it can be shown that $P\{E_h | \mathbf{H}\} = \omega^{-1} \sum_{k=1}^{\omega} P_{\text{bin}}(\lambda_k \tilde{\phi}_k \omega \bar{\gamma})$, where P_{bin} is the SER of the binary modulation and $\tilde{\phi}_k$ gives the power allocation for the header symbols. It follows that, for a given channel realization, the frame error rate can be expressed as

$$P_f = 1 - (1 - P\{E_h | \mathbf{H}\})^H (1 - P\{E_p | \mathbf{H}\})^{L/b}. \quad (36)$$

By considering (35) and (36), it can be seen that the P_f can be reduced by optimizing the power allocation for a given MIMO channel matrix realization. In effect, given $\bar{\gamma}$, the optimal values $\phi_1^*, \dots, \phi_\omega^*$ for a given realization of the fading coefficients $\lambda_1, \dots, \lambda_\omega$ are given by

$$(\phi_1^*, \dots, \phi_\omega^*) = \underset{\phi_1, \dots, \phi_\omega}{\operatorname{argmin}} \left\{ \frac{1}{\omega} \sum_{k=1}^{\omega} P_{M_k}(\lambda_k \phi_k \omega \bar{\gamma}) \right\} . \quad (37)$$

A solution for (37) is presented in Appendix C. Note that ϕ_k^* are functions of $\bar{\gamma}$ and of the fading coefficients; hence they are random variables. Using (35) and (37), it is found that, when the optimal power allocation is used, the mean SER is given by

$$\bar{P}\{E_p\}(\bar{\gamma}) = \frac{1}{\omega} \sum_{k=1}^{\omega} \mathbb{E} \{ P_{M_k}(\lambda_k \phi_k^* \omega \bar{\gamma}) \} . \quad (38)$$

Following the same line of thought, it is found that

$$\bar{P}\{E_h\}(\bar{\gamma}) = \frac{1}{\omega} \sum_{k=1}^{\omega} \mathbb{E} \left\{ P_{\text{bin}} \left(\lambda_k \tilde{\phi}_k^* \omega \bar{\gamma} \right) \right\} , \quad (39)$$

where $\tilde{\phi}_1^*, \dots, \tilde{\phi}_\omega^*$ are optimal values which are found by solving (37) for the specific case of a binary modulation.

Finally, using (30), (36), (38), (39), and assuming that the transmitter is equipped with a deep symbol interleaver [31], then the expected number of transmission trials for the proposed random feeding is obtained:

$$\bar{\tau}_\omega^N(\bar{\gamma}) = [1 - \bar{P}\{E_h\}(\bar{\gamma})]^{-H} [1 - \bar{P}\{E_p\}(\bar{\gamma})]^{-L/b} . \quad (40)$$

IV. OPTIMIZATION OF THE MIMO SCHEME

The goal of this section is to compare the performance of MIMO SVD, Beamforming[†], Zero Forcing and Generalized Alamouti schemes from an energy consumption point of view. For doing this, the optimization of physical layer parameters of each MIMO scheme (irradiation power, M -QAM constellation size and antenna array size) will be studied. Finally, the optimized performances of those MIMO schemes will be compared.

For simplicity, this Section focuses on the case $N_t = N_r = N$, although the generalization to the general case is straightforward.

[†]Note that MIMO Beamforming is just MIMO SVD when only the best angular direction is being used, i.e. $\omega = 1$.

A. Optimization of the mean irradiated power

The mean total energy consumption per goodbit (19) can be rewritten using (23), (30) and (40), and rewriting it so that the dependence on the average number of bits per symbol (b), on the multiplexing gain (ω) and on the antenna array size (N) becomes explicit. Concretely:

$$\bar{\mathcal{E}}_b = NS + \left[\hat{B}_N + NP_{\text{el}} + \omega Ad^\alpha \bar{\gamma} \right] \hat{T} \bar{\tau}_\omega^N(\bar{\gamma}) \quad (41)$$

$$= NS + \frac{1}{\omega R_s} \left(\hat{B}_N + NP_{\text{el}} + \omega Ad^\alpha \bar{\gamma} \right) \dots \\ \times \left(\frac{L+F}{bL} + \frac{H + \omega(NO_a + O_b)}{L} \right) \bar{\tau}_\omega^N(\bar{\gamma}) , \quad (42)$$

where $S = 2\mathcal{E}_{\text{st}}$ is the total startup energy consumption per branch and $P_{\text{el}} = P_{\text{el,tx}} + P_{\text{el,rx}}$ is the total electronic power consumption per transceiver branch (see Section II-A).

Using (41) it can be seen that the energy consumption is large at extreme values of the total SNR. In effect, if the SNR is low then the frame error rate tends to one, and hence the energy consumption is high because of the large number of retransmissions needed for a successful frame reception. On the contrary, at a high SNR (41) is also large because the radiated power (which is proportional to $\bar{\gamma}$) is excessive. We thus infer that an optimal SNR that minimizes the energy consumption must exist in between.

The previous analysis is analogous to the one made for single antenna systems in [20]. Following that work, let us define the SNR at which the system attains a minimal average energy consumption as

$$\bar{\gamma}^* = \underset{\bar{\gamma} \in [0, \infty)}{\operatorname{argmin}} \bar{\mathcal{E}}_b(\bar{\gamma}) . \quad (43)$$

Therefore, $\bar{\gamma}^*$ represents an optimal trade-off between radiation power and consumption because of retransmissions.

B. Optimization of MIMO SVD transmissions

Let us denote $\bar{\mathcal{E}}_{\text{SVD}}(M_1, \dots, M_\omega; \bar{\gamma}^*)$ as the mean energy consumption per goodbit of MIMO SVD communications when ω out of N eigenchannels are in use with M -ary modulations of sizes M_1, \dots, M_ω , and the optimal average SNR is being used. In this case, the baseband energy consumption is calculated using $\hat{\mathcal{E}}_{\text{cod,tx}} = \hat{\mathcal{E}}_{\text{cod,rx}} = \mathcal{E}_{\text{cod}}(\text{SVD})$, where $\mathcal{E}_{\text{cod}}(\text{SVD})$ is the average baseband energy consumption required for calculating one SVD of the MIMO channel matrix, which is the most demanding baseband operation in terms of energy consumption.

Consider optimizing the modulation sizes for a given number ω out of N eigenchannels in use over a link distance of d meters. Concretely, let us define

$$\bar{\mathcal{E}}_{\text{SVD}}^*(d, \omega, N) = \min_{\substack{M_k \in \mathcal{M} \\ (k=1, \dots, \omega)}} \bar{\mathcal{E}}_{\text{SVD}}(M_1, \dots, M_\omega; \bar{\gamma}^*) , \quad (44)$$

where \mathcal{M} is a set of modulations. Numerical evaluations of (44) were performed using typical low-power device parameters, which are shown in Table III, and $\mathcal{M} = \{2, 4, 16, 64\}$. For calculating \mathcal{E}_{SVD} it is used (12) with values as shown in Tables III and IV. Note that n_{add} , n_{prod} , n_{div} and n_{root} were determined based on the iterative algorithms presented in [35, Ch. 2.6]. Comparing the results for different values of ω , it is found that beamforming ($\omega = 1$), used along with BPSK, is the optimal choice for minimizing the energy consumption in long-range communications (see Figures 2 and 3). At the opposite end, minimum consumption for short transmission distances is attained using all the available eigenchannels ($\omega = N$) along with spectrally efficient modulations (large M).

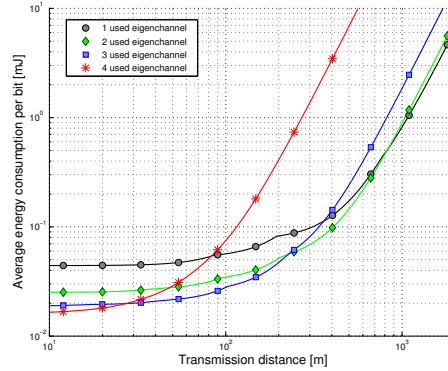


Fig. 2. Minimum energy consumption of 4x4 MIMO SVD communications using various numbers of used eigenchannels (ω). Beamforming ($\omega = 1$) is optimal for long-range communications. The optimal number of used eigenchannels increases as the link distance shortens.

[†]Source: [24]

[§]Source: IEEE 802.15.4 standard [36]

^{||}Source: [20]

^{*}Source: [37]

[‡]Source: [38]

^{††}Source: [39]

TABLE III
GENERIC LOW-POWER DEVICE PARAMETERS

Parameter	Description	Value
f_c	Carrier frequency	2.4 GHz [§]
v_{\max}	Maximal mobility speed	3 m/s
W	Bandwidth	10 kHz [†]
R_s	Symbol rate	10 kBaud [†]
L	Frame Payload	98 bytes [§]
O_a	Acquisition overhead	1 byte [§]
O_b	Estimation and synchronization overhead	3 bytes [§]
H	Frame header	26 bytes [§]
F	Feedback frame length	11 bytes [§]
\mathcal{E}_{st}	Start-up energy per branch	0.125 nJ [†]
α	Path-loss coefficient	3.2
A_0	Free space path loss	30 dB [†]
η	Power amplifier efficiency	35% [†]
$P_{\text{el,tx}}$	Tx electric power consumption	98.2 mW [†]
$P_{\text{el,rx}}$	Rx electric power consumption	112.5 mW [†]
N_0	Noise power density	-174 dBm/Hz
N_f	Receiver noise figure	10 dB [†]
M_l	Link margin	30 dB
f_{APU}	APU frequency	20 MHz [‡]
V_{dd}	APU voltage	3 V [‡]
I_0	Average current	6.37 mA [‡]
c_{add}	Adding cost	6 cycles [*]
c_{prod}	Product cost	13 cycles [*]
c_{div}	Division cost	21 cycles [*]
c_{root}	Root cost	149 cycles ^{††}

Both observations agree with intuition. At long link distances, the power consumed by the power amplifiers (\hat{P}_{PA}) dominates over the power consumed by the electronic components (\hat{P}_{el}). As there are no data rate constraints, it is therefore convenient to reduce the radiated power by investing it exclusively in the most favorable eigenchannel, and to choose BPSK for its low SNR requirement for attaining a given error rate. On the contrary, at short link distances the power consumed by electronic components dominates over the consumption of the power amplifiers.

TABLE IV
APU PARAMETERS FOR SVD

Parameter	Description	Value
n_{add}	APU cycles per addition	$\frac{16}{3}N^3 + 10N^2 - \frac{28}{3}N + 10$
n_{prod}	APU cycles per product	$\frac{16}{3}N^3 + 16N^2 - \frac{70}{3}N + 4$
n_{div}	APU cycles per division	$4N^2 - 2N - 3$
n_{root}	APU cycles per square root	$2N^2 - 3$

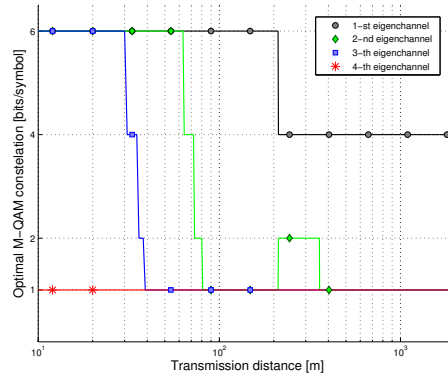


Fig. 3. Optimal size of the M -QAM constellation that each eigenchannel of a 4×4 MIMO system should use in order to minimize the overall energy consumption, when all the eigenchannels are in use (i.e. $\omega = 4$). Small modulations are optimal for long link distances, while the optimal modulation size grows as the link distance shortens.

The energy consumption (41) can be approximated for this case as

$$\bar{\mathcal{E}}_{\text{SVD}} \approx NS + [\hat{B}_N + NP_{\text{el}}]\hat{T} . \quad (45)$$

Hence, under these conditions, it is attractive to increase the throughput in order to reduce the total transmission time per bit (\hat{T}). This is attained by using all the eigenchannels along with large M -QAM constellations on those with good fading statistics (in the case shown in Figure 3, this includes all but the fourth eigenchannel, whose statistics are similar to a Rayleigh fading channel [29]) .

Let us define the minimal energy consumption per goodbit of an $N \times N$ MIMO SVD system

as

$$\bar{\mathcal{E}}_{\text{SVD}}^*(d, N) = \min_{\omega \in \{1, \dots, N\}} \bar{\mathcal{E}}_{\text{SVD}}^*(d, \omega, N) , \quad (46)$$

where $\bar{\mathcal{E}}_{\text{SVD}}^*(d, \omega, N)$ is given by (44). Numerical evaluations of (46) for different antenna array sizes (N) show that, for transmitting over long distances, large MIMO systems are more energy efficient than single-input single-output (SISO) systems (i.e. systems equipped with only one antenna, see Figure 4). This is because the large diversity degree of beamforming done with large antenna arrays generates important reductions in frame error rates. This, in turn, allows for reducing the radiated energy needed to reach the receiver with an adequate SNR. Large antenna arrays are however suboptimal for short transmission distances. This is because, as before, the reduction in the time per bit is not enough for compensating for the increase in electric power consumption ($N P_{\text{el}}$) and the larger overhead costs (c.f. (3)) demanded by larger MIMO configurations.

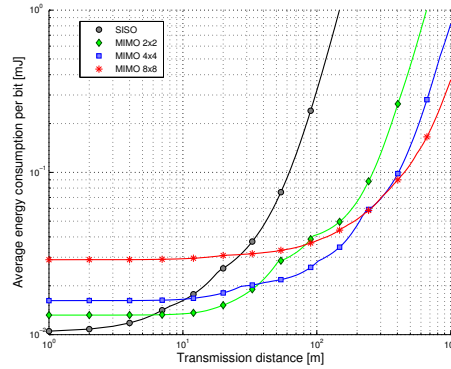


Fig. 4. Minimum energy consumption of different antenna array sizes over various link distances. Each MIMO system is using its own optimal number of used eigenchannels. Small systems consume less energy at short-range communications, while large antenna arrays achieve better performance when the link distance is large.

C. Performance of no-TCSI modulations

Let us denote $\bar{\mathcal{E}}_{\text{ZF}}(M; \bar{\gamma}^*)$ and $\bar{\mathcal{E}}_{\text{Al}}(M; \bar{\gamma}^*)$ as the mean energy consumption per goodbit of MIMO Zero-Forcing and Generalized Alamouti communications respectively, where the variables denote the size of the M -ary constellation and the optimal SNR $\bar{\gamma}^*(M)$. Note that the multiplexing gain of Zero-Forcing transmissions is $\omega_{\text{ZF}} = N_t$; for Generalized Alamouti modulations

the attainable multiplexing rate ω_A is given by [32]

$$\omega_A = \begin{cases} 1 & \text{for } N \in \{1, 2\}, \\ 3/4 & \text{for } N \in \{3, 4\}, \\ 1/2 & \text{for } N \geq 5. \end{cases} \quad (47)$$

As non-TCSI modulations involve simple precoding schemes, one can neglect the baseband energy consumption of the transmitter, $\hat{\mathcal{E}}_{\text{bb,tx}}$. The baseband energy consumption of the receiver, $\hat{\mathcal{E}}_{\text{bb,rx}}$, can also be neglected in the case of Generalized Alamouti transmissions, because the decoding consists only of few elementary arithmetic operations [32]. On the contrary, the consumption of Zero-Forcing receivers is non-negligible, as it involves the computation of the pseudo-inverse matrix of the MIMO channel $(\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger$ (see Appendix B-3). A popular way for calculating the pseudo-inverse involves the singular value decomposition of \mathbf{H} [40]. The computational cost of this method is dominated by the cost of computing the SVD, which is several order of magnitude larger than matrix-matrix multiplication [40]. Therefore, for Zero-Forcing receivers $\hat{\mathcal{E}}_{\text{bb,rx}}$ can be approximated by the cost of computing the SVD, which was considered in Section IV-B.

Define

$$\bar{\mathcal{E}}_{\text{ZF}}^*(d, N) = \min_{M \in \mathcal{M}} \bar{\mathcal{E}}_{\text{ZF}}(M; \bar{\gamma}) \quad (48)$$

$$\bar{\mathcal{E}}_{\text{AL}}^*(d, N) = \min_{M \in \mathcal{M}} \bar{\mathcal{E}}_{\text{AL}}(M; \bar{\gamma}) \quad (49)$$

as the energy consumption of Zero-Forcing and Generalized Alamouti modulations when the optimal modulation among a given set \mathcal{M} is used. Numerical evaluations of (48) and (49) were performed in the same way as the ones for MIMO SVD presented in Section IV-B (see Figures 5 and 6). It is found that Generalized Alamouti scheme, used along with BPSK, is the optimal choice for minimizing the energy consumption in long-range communications. This is because the large diversity gain of Generalized Alamouti scheme plays the same role as MIMO Beamforming in increasing the received SNR and improving the error rates. At the other end, Zero-Forcing used along with spectrally efficient modulations (large M) is optimal for transmissions over short link distances. This is because the multiplexing gain allows for reducing the total air time per bit, in the same fashion as when all the eigenchannels of the SVD modulation are used.

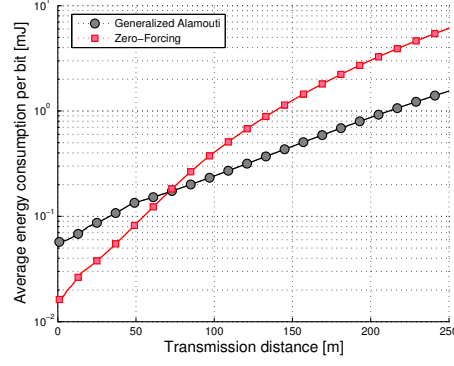


Fig. 5. Minimum energy consumption of 4x4 MIMO communications when Zero-Forcing or Generalized Alamouti modulations are used. Generalized Alamouti modulations, which provide a large diversity gain, are the optimal strategy for long-range communications. On the contrary, Zero-Forcing modulation, which provides a large multiplexing gain, is energy-optimal when the link distance is short.

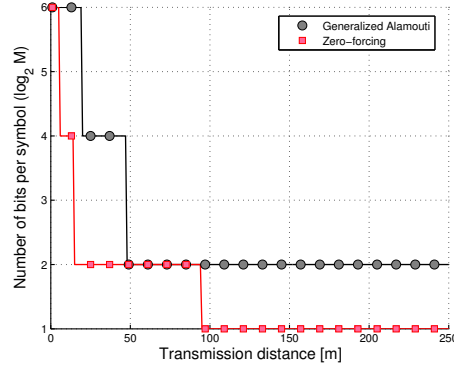


Fig. 6. Size of the M -QAM modulation that minimizes the energy consumption of a 4×4 MIMO system over various link distances for Zero-Forcing and Generalized Alamouti modulations. Small modulations are optimal for long link distances, while the optimal modulation size grows as the link distance shortens. In contrast to Figure 4, in this case SISO systems are optimal for links of less than 60 meters.

Let us characterize the optimal performance of a system of N antennas in the considered no-TCSI scenario as

$$\bar{\mathcal{E}}_{\text{noTCSI}}^*(d, N) = \min \left\{ \bar{\mathcal{E}}_{\text{ZF}}^*(d, N), \bar{\mathcal{E}}_{\text{AL}}^*(d, N) \right\} . \quad (50)$$

Numerical evaluations show that the optimal antenna array size depends on the link distance in similar way to what was found for the MIMO SVD case. In effect, large MIMO systems

are more energy efficient for transmitting over long link distances (see Figure 7), because the diversity degree of Generalized Alamouti schemes over large antenna arrays allows for reducing the radiated energy that is necessary for reaching the receiver with an adequate SNR. However, large antenna arrays are suboptimal for short-range communications, because the reduction in the time per bit gained by the Zero-Forcing multiplexing gain does not compensate for the higher power consumption introduced by larger antenna arrays.

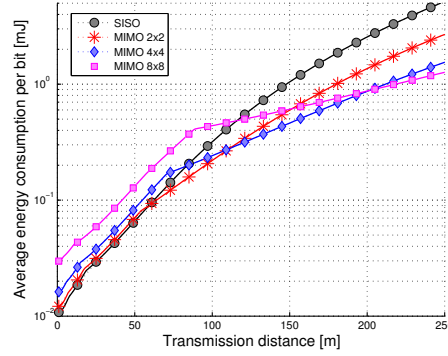


Fig. 7. Minimum energy consumption of different antenna array sizes when the best choice between ZF and Generalized Alamouti schemes is used. Small systems generally consume less energy in short-range transmissions, while large antenna arrays achieve a better performance when the link distance is large.

Let us define the relative energy loss for not having TCSI as

$$\mathcal{L}(d) = \frac{\min_N \bar{\mathcal{E}}_{\text{noTCSI}}^*(d, N)}{\min_N \bar{\mathcal{E}}_{\text{SVD}}^*(d, N)} . \quad (51)$$

For all reasonable link distances \mathcal{L} is greater than 1, which means that MIMO SVD used with the adequate antenna array size is always more energy efficient than non-TCSI modulations (see Figure 8). The relative energy loss is approximately linear to the transmission distance, being very significative for long-range communications. Nevertheless, the energy cost of not exploiting TCSI diminishes as the transmission distance shortens, and becomes negligible for short-range communications. This would become more pronounced if channels with a line-of-sight component are considered, as the evaluations have been done using Rayleigh fading statistics –which is the most unfavorable scenario from the SNR point of view.

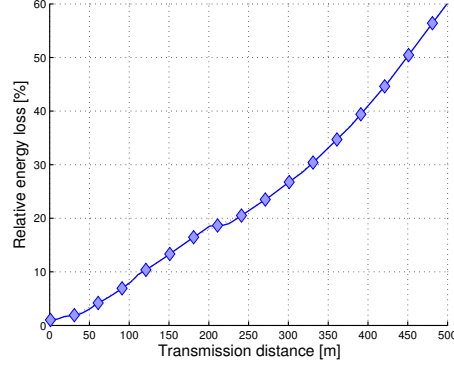


Fig. 8. Relative energy loss for not having TCSI in the studied scenario, which is calculated as the ratio between the minimal energy consumption of ZF and Generalized Alamouti MIMO schemes, which do not exploit TCSI, and MIMO SVD.

V. CONCLUSIONS

We studied the energy consumption for delivering one bit of data without error over a MIMO communication link. For a given size of the antenna array, we show that energy-efficient communications over long link distances is achieved by schemes with low throughput – i.e. large diversity gain and small symbol constellations. This is because the energy consumption is dominated by the radiated energy, and these schemes reduce the amount of irradiated power needed to achieve a low symbol error rate. On the contrary, schemes with high throughput are optimal for performing transmissions over short link distances. In this case, the energy budget is dominated by the power consumption of electronic components. Hence, schemes with large multiplexing gain and large symbol constellations are optimal because of they reduce the average time per bit.

We show that large MIMO systems can achieve significant savings when the link distance is large. This is a consequence of the lower SNR requirement for achieving a small symbol error rate of MIMO schemes with good diversity gain. On the contrary, small antenna arrays outperform large systems in performing short range transmissions. This is because the smaller time per bit achieved by the multiplexing capabilities of large MIMO systems do not compensate the increase in electronic power and transmission overhead demanded by the multiple-antenna array. In this respect, interesting analogies between the analysed schemes were found: Generalized Alamouti perform similarly to MIMO Beamforming because of their high diversity gain and

low throughput, while Zero Forcing receivers are similar to SVD with power loaded over all the eigenchannels because of their spectral efficiency and unfavorable error statistics.

Finally, we found that MIMO SVD is always more energy-efficient than ZF or Generalized Alamouti schemes. Nevertheless, although the savings can be larger than 60 times in long-range transmissions, they diminish when the transmission distance shortens and become irrelevant when the link distance is small. These results suggest that exploiting TCSI is not critical for performing energy-efficient short-range communications.

APPENDIX A

DISTRIBUTION OF τ

For finding the distribution of the number of transmission trials until a frame is decoded without error, τ , let us first consider a random variable X_t whose value is 1 if the t -th transmission trial is decoded successfully and zero otherwise. One could picture X_t as a coin whose fairness is also a random variable. Then, using the fact that the conditional probability is, by definition, calculated by $\mathbb{P}\{X_t = 0|P_f(t)\} = P_f(t)$ (where $\mathbb{P}\{A\}$ denotes the probability of the event A), then it is observed that

$$\mathbb{P}\{X_t = 0\} = \int_0^\infty \mathbb{P}\{X_t = 0|P_f\} f_{P_f}(P_f) dP_f \quad (52)$$

$$= \int_0^\infty P_f f_{P_f}(P_f) dP_f = \bar{P}_f, \quad (53)$$

where $f_{P_f}(P_f)$ is the p.d.f. of the t -th frame error rate (the index t has been dropped for simplicity). Note that the first equality is obtained using Bayes rule. This result shows that X_t is a Bernoulli random variable with parameter \bar{P}_f .

The distribution of τ can be found now by a direct calculation:

$$\mathbb{P}\{\tau = r\} = \mathbb{P}\{X_1 = 0, \dots, X_{r-1} = 0, X_r = 1\} \quad (54)$$

$$= \mathbb{P}\{X_r = 1\} \prod_{t=1}^{r-1} \mathbb{P}\{X_t = 0\} \quad (55)$$

$$= (1 - \bar{P}_f) \bar{P}_f^{r-1}, \quad (56)$$

which gives the desired result.

APPENDIX B

ERROR STATISTICS OF VARIOUS MIMO SCHEMES

For completeness, we present the derivation of the mean symbol error rate of MIMO SVD, Zero-Forcing and Generalized Alamouti schemes. We focus on the case of a narrowband MIMO channel, in which the received signal vector $\mathbf{v} = (v_1, \dots, v_{N_r})^t$ can be expressed in terms of the transmitted signal vector $\mathbf{u} = (u_1, \dots, u_{N_t})^t$ as

$$\mathbf{v} = \mathbf{H}\mathbf{u} + \mathbf{w} \quad , \quad (57)$$

where u_j is the complex symbol transmitted through the j -th antenna, v_i is the complex symbol received by the i -th transmission branch, \mathbf{H} is a $N_r \times N_t$ random matrix with coefficients $h_{i,j}$ which are i.i.d. standard complex normal random variables and $\mathbf{w} = (w_1, \dots, w_n)^t$ is the vector of additive white gaussian noise terms experimented in each branch of the receiver, which satisfy $\mathbb{E}\{\mathbf{w}\mathbf{w}^H\} = \sigma_n^2 \mathbf{I}_{N_r \times N_r}$ where σ_n is the noise power and $\mathbf{I}_{N_r \times N_r}$ is the $N_r \times N_r$ identity matrix [17].

1) *MIMO SVD*: Consider the case in which the transmitter is sending data over ω out of the $N = \min\{N_t, N_r\}$ available eigenchannels feeding the SVD encoder in a pseudo-random fashion, as defined in III-B2. Define A_k as the event in which a symbol is assigned to the k -th eigenchannel and E_p as the event that a payload symbol was decoded with error. As all events A_k are jointly exhaustive and mutually exclusive, one can decompose the symbol-error events as $E_p = \cup_{k=1}^{\omega} (E_p \cap A_k)$. Hence, for a given realization of the channel matrix \mathbf{H} , the average error rate of a payload symbol can be written in terms of the M -ary modulation symbol error rate, $P_M(\gamma)$, as

$$P\{E_p|\mathbf{H}\} = \mathbb{P}\{\cup_{k=1}^{\omega} (E_p \cap A_k)|\mathbf{H}\} \quad (58)$$

$$= \sum_{k=1}^{\omega} \mathbb{P}\{E_p \cap A_k|\mathbf{H}\} \quad (59)$$

$$= \sum_{k=1}^{\omega} \mathbb{P}\{A_k\} \mathbb{P}\{E_p|A_k, \mathbf{H}\} \quad (60)$$

$$= \frac{1}{\omega} \sum_{k=1}^{\omega} P_{M_k}(\text{SNR}_k) \quad , \quad (61)$$

where M_k denotes the modulation that is used among the k -th eigenchannel. Above, it is assumed that $\mathbb{P}\{A_k\} = 1/\omega$ for all $k = 1 \dots \omega$.

2) *Generalized Alamouti schemes:* Following [32], let us assume that the complex symbol s is transmitted through the MIMO link using D different beam vectors \mathbf{x}_i successively. According to (57), the information about s available to the receiver will be contained in D received vectors \mathbf{y}_i in the form of

$$\mathbf{y}_i = \mathbf{H}\mathbf{x}_i s + \mathbf{w}_i . \quad (62)$$

If linear receivers are used, then the decoding decision will be taken on a statistic given by

$$Z = \sum_{i=1}^D \left(\mathbf{a}_i^\dagger \mathbf{H}\mathbf{x}_i s + \mathbf{a}_i^\dagger \mathbf{w}_i \right) , \quad (63)$$

where \mathbf{a} are arbitrary post-processing complex vectors. It can be shown that the highest achievable SNR, which is attained when $\mathbf{a}_i = \mathbf{H}\mathbf{x}_i$, is given by

$$\text{SNR}^* = \sum_{i=1}^D \|\mathbf{H}\mathbf{x}_i\|^2 \frac{E_s}{\sigma_n^2} = \text{tr} \{ \mathbf{H}\mathbf{R}_x \mathbf{H}^\dagger \} \frac{E_s}{\sigma_n^2} , \quad (64)$$

where $\mathbb{E}\{|s|^2\} = E_s$ is the mean symbol power, $\text{tr}\{\cdot\}$ is the trace operator and $\mathbf{R}_x = \sum_{i=1}^D \mathbf{x}_i \mathbf{x}_i^\dagger$ is the $N_t \times N_t$ correlation matrix of the beam vectors. In order to preserve the energy per symbol $\sum_{i=1}^D \|\mathbf{x}_i\|^2 = 1$ is required, where $\|\cdot\|$ is the euclidean norm. A direct calculation shows that $\mathbf{R}_x = \mathbf{X}\mathbf{X}^\dagger$, with \mathbf{X} the $N_t \times D$ the matrix which columns are the beam vectors \mathbf{x}_i . Therefore, \mathbf{R}_x is a Gramian matrix and therefore all its eigenvalues r_i must be real and non-negative [41].

It can also be shown that

$$\sum_{j=1}^D r_i = \text{tr} \{ \mathbf{R}_x \} = \text{tr} \{ \mathbf{X}\mathbf{X}^\dagger \} = \sum_{i=1}^D \sum_{j=1}^{N_t} \|x_{i,j}\|^2 = 1 , \quad (65)$$

where the last equality comes from the normalization condition.

We want to maximize (64) under the assumption that the transmitter has no CSI, which implies that \mathbf{R}_x has to be determined independently from the channel matrix \mathbf{H} . When CSI is available in the transmitter, the optimal choice is given by $\mathbf{R}_x = \mathbf{v}_{\max} \mathbf{v}_{\max}^\dagger$, where \mathbf{v}_{\max} is the first right singular vector of the MIMO channel matrix \mathbf{H} which points out the most favorable angular direction [17]. This is a rank-one matrix with $r_1 = 1$ and $r_i = 0$ for all $i = 2 \dots D$, which points out the fact that there is one preferred direction. When there is no CSI available there is no a priori preferred directions [17], and therefore all the eigenvalues should be equal. Therefore, if $D \leq N_t$, the correlation matrix is given by (up to unitary transformations, i.e. change of basis)

$$\mathbf{R}_x = \frac{1}{D} \mathbf{I}_D^{N_t} = \frac{1}{D} \left(\begin{array}{c|c} \mathbf{I}_{D \times D} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0}_{N_t-D \times N_t-D} \end{array} \right) \quad (66)$$

where $\mathbf{I}_D^{N_t}$ is a $N_t \times N_t$ diagonal matrix with D ones and $N_t - D$ zero diagonal terms. Using this choice of \mathbf{R}_x on (64), the optimal SNR under the no TCSI assumption can be found to be

$$\text{SNR}_{\text{noTCSI}}^* = \frac{\text{tr}\{\mathbf{H}\mathbf{I}_D^{N_t}\mathbf{H}^\dagger\}}{D} \frac{E_s}{\sigma_n^2} = \frac{1}{D} \sum_{j=1}^{N_r} \sum_{k=1}^D |h_{j,k}|^2 \frac{E_s}{\sigma_n^2} . \quad (67)$$

As each term $h_{j,k}$ is an independent standard complex gaussian random variable, (67) is a sum of DN_r independent exponential random variables [42] and $\mathbb{E}\{\text{SNR}_{\text{noTCSI}}^*\} = N_r E_s / \sigma_n^2$. Therefore

$$\text{SNR}_{\text{noTCSI}}^* \sim \Gamma\left(DN_r, N_r \frac{E_s}{\sigma_n^2}\right) \quad (68)$$

is a gamma random variable, which is equivalent to say that the error statistics are the same than in a Nakagami- m fading channel with power gain $\mu = N_r$ and diversity degree $m = DN_r$ [34]. It can be seen from (66) and (68) that the largest diversity gain is achieved when $D = N_t$, while no further improvements can be obtained using larger values of D . This agrees with the intuitive fact that maximum diversity gain is achieved when the data symbols are sent using a maximal linear independent set of N_t -dimensional beam vectors \mathbf{x}_i .

3) *Multiplexing*: Zero-Forcing receiver is a linear decorrelator of the MIMO channel matrix [17]. The post-processed signal of a Zero-Forcing receiver is given by

$$\mathbf{r} = (\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{y} = \mathbf{x} + (\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{w} . \quad (69)$$

It can be shown that, under the present assumptions, the MIMO channel matrix \mathbf{H} have full rank almost surely [42], and therefore $\mathbf{H}^\dagger \mathbf{H}$ is invertible if $N_r \geq N_t$.

Following [33], let us look for the SNR of the k -th received symbol r_k . The correlation matrix of the colored noise $\mathbf{n} = (\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{w}$ can be found to be

$$\mathbf{R}_n = \mathbb{E}\{(\mathbf{H}^\dagger \mathbf{H})^{-1} \mathbf{H}^\dagger \mathbf{w} \mathbf{w}^\dagger \mathbf{H} (\mathbf{H}^\dagger \mathbf{H})^{-1}\} \quad (70)$$

$$= \sigma_n^2 (\mathbf{H}^\dagger \mathbf{H})^{-1} . \quad (71)$$

Hence, the noise power that influences the decoding of r_k can be calculated as

$$\sigma^2(k) = [\mathbf{R}_n]_{kk} = \sigma_n^2 [(\mathbf{H}^\dagger \mathbf{H})^{-1}]_{k,k} \quad (72)$$

where $[\cdot]_{k,k}$ denotes the coefficient with row k and column k of a given matrix. Using (72), the SNR of the k -th symbol is given by

$$\text{SNR}_k = \frac{E_s}{\sigma^2(k)} = \frac{1}{[(\mathbf{H}^\dagger \mathbf{H})^{-1}]_{k,k}} \frac{E_s}{\sigma_n^2} . \quad (73)$$

It is known that $\mathbf{H}^\dagger \mathbf{H}$ distributes as a complex Wishart random matrix [43]. Furthermore, it can be shown that for any complex Wishart random matrix $\mathbf{A} \sim \tilde{W}_M(N, \mathbf{I})$ with $N \geq M$ and any vector \mathbf{y} then [42]

$$\frac{\|\mathbf{y}\|^2}{\mathbf{y}^* \mathbf{A}^{-1} \mathbf{y}} \sim \Gamma(N - M + 1, 1) , \quad (74)$$

where $\Gamma(m, \mu)$ is a Gamma random variable with p.d.f. given by (32). Using $\mathbf{y} = \mathbf{e}_k$ (the all-zeros vector except a one in the k -th component) in (73) it can be seen that

$$\text{SNR}_k \sim \Gamma\left(N_r - N_t + 1, \frac{E_s}{\sigma_n^2}\right) , \quad (75)$$

which is equivalent to say that (69) is equivalent to N_r Nakagami- m channels with diversity gain $m = N_r - N_t + 1$ [29].

APPENDIX C

SOLUTION OF (37)

The problem stated in (37) is equivalent to the following optimization problem:

$$\begin{aligned} & \text{minimize} \quad f(\phi_1, \dots, \phi_\omega) = \frac{1}{\omega} \sum_{k=1}^{\omega} P_{M_k}(\lambda_k \phi_k \omega \bar{\gamma}) \\ & \text{subject to} \quad \begin{cases} \sum_{k=1}^{\omega} \phi_k = 1, \\ \phi_k \geq 0 \quad \text{for all } k = 1, \dots, \omega. \end{cases} \end{aligned} \quad (76)$$

Above, the SER function is given by

$$P_{M_k}(\lambda_k \phi_k \omega \bar{\gamma}) = \begin{cases} c_k Q(\sqrt{a_k \lambda_k \phi_k \omega \bar{\gamma}}) & \text{for BPSK} \\ 1 - [1 - c_k Q(\sqrt{a_k \lambda_k \phi_k \omega \bar{\gamma}})]^2 & \text{for } M\text{-QAM} \end{cases} \quad (77)$$

where $Q(x)$ is the tail probability of the standard normal distribution and a_k and c_k are appropriate constants [31]. We will first show that the previous optimization problem is convex and then apply the Karush-Kuhn-Tucker (KKT) conditions.

Let us write the objective function as $f = \sum_{k=1}^{\omega} f_k / \omega$, where $f_k = P_{M_k}(\lambda_k \phi_k \omega \bar{\gamma})$ as given by (77). It can be shown that $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ is a convex function for any value of k , because it only depends on ϕ_k and $\frac{\partial^2 f_k}{\partial \phi_k^2} > 0$. Therefore, f is also convex, since it is the sum of convex functions. The domain of allowable solutions is a probability simplex, which is a well known convex set [44].

As the above problem is convex, the KKT conditions are necessary and sufficient to characterize the optimal solution [44]. Therefore, expressions for the optimal values $(\phi_1^*, \dots, \phi_\omega^*)$ can be found using the following Lagrangian:

$$\Lambda = \frac{1}{\omega} \sum_{k=1}^{\omega} P_{M_k}(\lambda_k \phi_k \omega \bar{\gamma}) + \chi_0 \left(\sum_{k=1}^{\omega} \phi_k - 1 \right) - \sum_{k=1}^{\omega} \chi_k \phi_k . \quad (78)$$

where χ_k are Lagrange multipliers associated to the restrictions of the problem. Using the KKT conditions on (78), it can be shown that the optimal values ϕ_k^* satisfy the conditions

$$\frac{s_k \phi_k^* e^{s_k \phi_k^*}}{[1 - r_k Q(\sqrt{s_k \phi_k^*})]^2} = \beta \lambda_k^2 \quad \forall k \in \{1, \dots, n\} , \quad (79)$$

where $s_k = a \lambda_k \omega \bar{\gamma}$, β is a constant, and $r_k = 2(1 - 1/\sqrt{M_k})$ for M -QAM modulations or $r = 0$ for BPSK.

Solving (79) for BPSK can be done using the Lambert function $W(x)$ [45]. In order to solve (79) for the general M -QAM case, we define a generalized Lambert function $W_r(x)$ as

$$\frac{W_r(x) e^{W_r(x)}}{[1 - r Q(\sqrt{W_r(x)})]^2} = x . \quad (80)$$

The function $W_r(x)$ is well defined for all $x \geq 0$ and $r \geq 0$, because it is the inverse of a strictly increasing function.

Using the generalized Lambert function, the solution to (76) can be written as

$$\phi_k^* = \frac{1}{a \lambda_k \omega \bar{\gamma}} W_{r_k}(\beta \lambda_k^2) . \quad (81)$$

The constant β is the number that satisfies the condition $\sum_{k=1}^{\omega} \phi_k^* = 1$. By replacing (81) into this condition we find

$$a \omega \bar{\gamma} = \sum_{k=1}^{\omega} \frac{1}{\lambda_k} W_{r_k}(\beta \lambda_k^2) , \quad (82)$$

from where β must be found numerically.

Is to be noted that (80) can be used to build a look-up table for W_r . Using such a table and the derivative of W_r (which can be calculated from (80) by implicit derivation), a Newton-Raphson algorithm can be built for finding β from (82). This can be achieved by a small average number of iterations, and therefore requires a negligible amount of energy. Nevertheless, a specific discussion on implementation issues goes out of the scope of the present work.

REFERENCES

- [1] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell labs technical journal*, vol. 1, no. 2, pp. 41–59, 1996.
- [2] E. T. Ar and I. E. Telatar, "Capacity of multi-antenna gaussian channels," *European Transactions on Telecommunications*, vol. 10, pp. 585–595, 1999.
- [3] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.
- [4] G. Bravos and A. G. Kanatas, "Energy efficiency comparison of MIMO-based and multihop sensor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 23, pp. 1–13, 2008.
- [5] M. Z. Siam, M. Krunz, S. Cui, and A. Muqattash, "Energy-efficient protocols for wireless networks with adaptive MIMO capabilities," *Wireless Networks Journal*, vol. 16, no. 1, pp. 199–212, Jan. 2010.
- [6] E. V. Belmega and S. Lasaulce, "An information-theoretic look at MIMO energy-efficient communications," in *Proc. of VALUETOOLS '09*, vol. 58. Brussels, Belgium: ICST, 2009, pp. 1–10.
- [7] H. Kim, C.-B. Chae, G. de Veciana, and R. Heath, "A cross-layer approach to energy efficiency for adaptive MIMO systems exploiting spare capacity," *IEEE Transactions on Wireless Communications*, vol. 8, no. 8, pp. 4264–4275, Aug. 2009.
- [8] R. Prabhu and B. Daneshrad, "Energy-Efficient Power Loading for a MIMO-SVD System and Its Performance in Flat Fading," in *Proc. GLOBECOM 2010*, Dec. 2010, pp. 1–5.
- [9] F. Heliot, M. Imran, and R. Tafazolli, "On the energy efficiency gain of MIMO communication under various power consumption models," in *Proc. FutureNet 2011*, June 2011, pp. 1–9.
- [10] C. Jiang and L. Cimini, "Energy-efficient multiuser MIMO beamforming," in *Proc. CISS 2011*, Mar. 2011, pp. 1–5.
- [11] C.-L. Chen, W. Stark, and S.-G. Chen, "Energy-Bandwidth Efficiency Tradeoff in MIMO Multi-Hop Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1537–1546, September 2011.
- [12] J. Xu and L. Qiu, "Energy Efficiency Optimization for MIMO Broadcast Channels," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 690–701, February 2013.
- [13] H. Li, L. Song, D. Zhu, and M. Lei, "Energy efficiency of large scale MIMO systems with transmit antenna selection," in *Proc. IEEE ICC 2013*, June 2013, pp. 4641–4645.
- [14] J. Jiang, M. Dianati, M. A. Imran, and R. Tafazolli, "On the Energy Efficiency of MIMO Channels in Correlated Rayleigh Fading Environment," in *Proc. of the 19th European Wireless Conference*, April 2013, pp. 1–5.
- [15] J. Chen, F. Talebi, and T. Pratt, "Energy efficiency of co-polarized and space-polarization mimo architectures in packet-based communication systems," in *Proc. of IEEE MILCOM 2013*, Nov 2013, pp. 311–316.
- [16] G. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Wireless Communications*, vol. 18, no. 6, pp. 28–35, Dec. 2011.
- [17] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [18] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.
- [19] F. Rosas and C. Oberli, "Energy-efficient MIMO SVD communications," in *Proc. IEEE PIMRC'12*, Sep. 2012.
- [20] —, "Modulation and SNR optimization for achieving energy-efficient communications over short-range fading channels," *IEEE Transactions on Wireless Communications*, vol. 11, no. 12, pp. 4286–4295, 2012.

- [21] D. Bates, S. Henriksen, B. Ninness, and S. R. Weller, "A 4x4 FPGA-based wireless testbed for LTE applications," in *Proc. PIMRC 2008*, 2008, pp. 1–5.
- [22] K. Nishimori, R. Kudo, N. Honma, Y. Takatori, and M. Mizoguchi, "16x16 Multiuser MIMO Testbed Employing Simple Adaptive Modulation Scheme," in *Proc. VTC Spring 2009*, Apr. 2009, pp. 1–5.
- [23] Mango Communications, "WARP MIMO Kit," 2013. [Online]. Available: <http://mangocomm.com/>
- [24] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, 2005.
- [25] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. John Wiley & Sons, Inc., 1998.
- [26] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*. Cambridge University Press, 1998.
- [27] A. Sinha and A. Chandrakasan, "Dynamic power management in wireless sensor networks," *IEEE Design & Test of Computers*, vol. 18, no. 2, pp. 62–74, 2001.
- [28] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ, USA: Prentice Hall, 2002.
- [29] F. Rosas and C. Oberli, "Nakagami- m approximations for multiple-input multiple-output singular value decomposition," *IET Transactions on Communications*, vol. 7, no. 6, pp. 554–561, April 2013.
- [30] R. Gray, *Probability, Random Processes, and Ergodic Properties*. Springer-Verlag, 1988.
- [31] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
- [32] G. Ganesan and P. Stoica, "Space-time block codes: a maximum SNR approach," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1650–1656, May 2001.
- [33] D. Gore, J. Heath R.W., and A. Paulraj, "On performance of the zero forcing receiver in presence of transmit correlation," in *Proc. of IEEE International Symposium on Information Theory*, 2002, p. 159.
- [34] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels*. John Wiley & Sons Inc., 2000.
- [35] A. Björck, *Numerical Methods for Least Squares Problems*. Philadelphia: SIAM, 1996.
- [36] *Specifications for Local and Metropolitan Area Networks- Specific Requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY)*, IEEE Std. 802.15.4-2006, 2006.
- [37] T. Instruments, "MSP430x5xx/MSP430x6xx Family User's Guide," Online, 2010.
- [38] —, "MSP430BT5190 Product Review," Online, 2010.
- [39] R. M. Fosler, "Fast Integer Square Root," Microchip Technology Inc., 2000.
- [40] G. H. Golub and C. F. V. Loan, *Matrix Computations*. The Johns Hopkins University Press; 3rd edition, 1996.
- [41] P. Lax, *Linear Algebra and Its Applications*, ser. Pure and Applied Mathematics. Wiley, 2007, no. v. 10.
- [42] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. Hoboken, NJ: John Wiley & Sons, Ltd., 1982.
- [43] A. Edelman, "Eigenvalues and condition numbers of random matrices," Ph.D. dissertation, Department of Mathematics, MIT, 1989.
- [44] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.
- [45] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Advances in Computational Mathematics*, vol. 5, pp. 329–359, 1996.